Assignment

Q1. Load the "titanic" dataset using the load_dataset function of seaborn. Use Plotly express to plot a
scatter plot for age and fare columns in the titanic dataset.

Ans:To create a scatter plot for the "age" and "fare" columns in the Titanic dataset using Plotly Express, you can follow the code below. First, make sure to install the Plotly library if you haven't already:

bash

Copy code

Now, you can use the following Python code:

python

Copy code

```
import         as
import                 as


                                "titanic"


                                          'age'    'fare'        'Scatter Plot of
Age and Fare in Titanic Dataset'
```

In this code:

- `sns.load_dataset("titanic")` loads the "titanic" dataset using Seaborn.
- `px.scatter()` creates a scatter plot using Plotly Express, specifying 'age' for the x-axis and 'fare' for the y-axis.
- `title` parameter is used to add a title to the plot.
- `show()` displays the plot.

This will open a new browser window or notebook cell displaying the interactive scatter plot created with Plotly Express. You can zoom, pan, and hover over points to explore the data interactively. Adjust the code according to your preferences and analysis requirements.

Q2. Using the tips dataset in the Plotly library, plot a box plot using Plotly express.
Ans:Plotly Express makes it easy to create a box plot for the "tips" dataset. Here's an example:
python

Copy code

```
import                as

                          'day'    'total_bill'        'sex'          True
    'Box Plot of Total Bill by Day and Gender'
```

In this code:

- `px.data.tips()` loads the "tips" dataset from Plotly Express.
- `px.box()` creates a box plot, specifying 'day' for the x-axis, 'total_bill' for the y-axis, and 'sex' for the color parameter.
- `notched=True` adds notches to the box plots for visualizing the confidence interval around the median.
- `title` parameter is used to add a title to the plot.
- `show()` displays the interactive box plot.

This code generates an interactive box plot for the "total_bill" column grouped by "day" and colored by "sex" from the "tips" dataset. You can customize the parameters according to your preferences and analysis needs.

Q3. Using the tips dataset in the Plotly library, Plot a histogram for x= "sex" and y="total_bill" column in
the tips dataset. Also, use the "smoker" column with the pattern_shape parameter and the "day" column with the color parameter.

Ans:can create a histogram for the "sex" and "total_bill" columns in the "tips" dataset using Plotly Express, and customize the appearance using the "smoker" and "day" columns. Here's an example:
python

Copy code

```
import                as


                                    'sex'    'total_bill'        'day'
          'smoker'
               'smoker'        'Histogram of Total Bill by Sex and Smoking
Status'
```

In this code:

- `px.data.tips()` loads the "tips" dataset from Plotly Express.
- `px.histogram()` creates a histogram, specifying 'sex' for the x-axis, 'total_bill' for the y-axis, and using 'day' for color and 'smoker' for both pattern_shape and facet_col.
- `title` parameter is used to add a title to the plot.
- `show()` displays the interactive histogram.

This code generates an interactive histogram for the "total_bill" column, with "sex" on the x-axis, colored by "day," and using different shapes for "smoker" categories. The facet_col parameter separates the data into facets based on the "smoker" column. You can adjust the parameters based on your preferences and analysis needs.

Q4. Using the iris dataset in the Plotly library, Plot a scatter matrix plot, using the "species" column for
the color parameter.
Ans:To create a scatter matrix plot for the "iris" dataset using Plotly Express and color the points based on the "species" column, you can use the following code:
python

Copy code

```
import                as


                                                        'sepal_width'
'sepal_length'  'petal_width'  'petal_length'
       'species'        'Scatter Matrix Plot of Iris Dataset'
```

In this code:

- `px.data.iris()` loads the "iris" dataset from Plotly Express.
- `px.scatter_matrix()` creates a scatter matrix plot, specifying the dimensions as the columns ['sepal_width', 'sepal_length', 'petal_width', 'petal_length'] and using the "species" column for color.
- `title` parameter is used to add a title to the plot.
- `show()` displays the interactive scatter matrix plot.

This code generates an interactive scatter matrix plot for the "iris" dataset, where each point is colored based on the "species" column. You can adjust the parameters based on your preferences and analysis needs.

Note: Use "sepal_length", "sepal_width", "petal_length", "petal_width" columns only with the dimensions parameter.
Q5. What is Distplot? Using Plotly express, plot a distplot.
Ans:A distplot (distribution plot) is a type of plot that combines a histogram with a kernel density estimate (KDE) and potentially rug plot elements. It provides a visual representation of the distribution of a univariate dataset.

To create a distplot using Plotly Express, you can use the `px.histogram()` function with the `marginal` parameter set to "rug" for the rug plot and the `marginal_y` parameter set to "box" for a box plot along the y-axis. Here's an example:

python

Copy code

```
import                as


                                  'sepal_length'          'rug'
        'box'
      'Distplot of Sepal Length in Iris Dataset'
```

In this code:

- `px.data.iris()` loads the "iris" dataset from Plotly Express.
- `px.histogram()` creates a histogram for the "sepal_length" column, and the `marginal` and `marginal_y` parameters are used to add a rug plot along the x-axis and a box plot along the y-axis.
- `title` parameter is used to add a title to the plot.
- `show()` displays the interactive distplot.

This code generates an interactive distplot for the "sepal_length" column in the "iris" dataset, providing a visual representation of the distribution along with a rug plot and a box plot. Adjust the parameters based on your preferences and analysis needs.